Heat wave attribution over Europe

Kamal Gasser [1] Francesco Ragone [2] Johan Segers [1,3]

[1]LIDAM, ISBA
UCLouvain, Belgium

[2]School of Computing and Mathematical Sciences
University of Leicester, UK

[3]Department of Mathematics
KU Leuven, Belgium

Introduction

- Research question:
  To what extent does anthropogenic forcing increase the frequency, intensity, and duration of extreme heat events?
- Method: Compare CMIP6 climate simulations with and without anthropogenic forcing using space-time extreme value theory.
- What do we focus on?
  Quantifying changes in heat wave characteristics over Europe.

Datasets

- Observational data: Daily maximum temperature from the Berkeley Earth dataset [2] ($1°\times1°$ resolution, 1880–present), combining over 30 000 weather stations with spatial reconstruction for exploratory analysis and clustering.
- Model simulations: CMIP6 climate models under **historical** (with human forcing, Factual) and **natural** (without human forcing, Counterfactual) runs (1850–2020).

Clustering region

- How: Partitioning around medoids.
- Distance based on pairwise tail dependence coefficient [3,4]:

$$d(\mathbf{s}_l, \mathbf{s}_k \mid u) = \frac{1-\chi_{\mathbf{s}_l,\mathbf{s}_k,t_q,t_r}(u)}{2(3-\chi_{\mathbf{s}_l,\mathbf{s}_k,t_q,t_r}(u))}$$

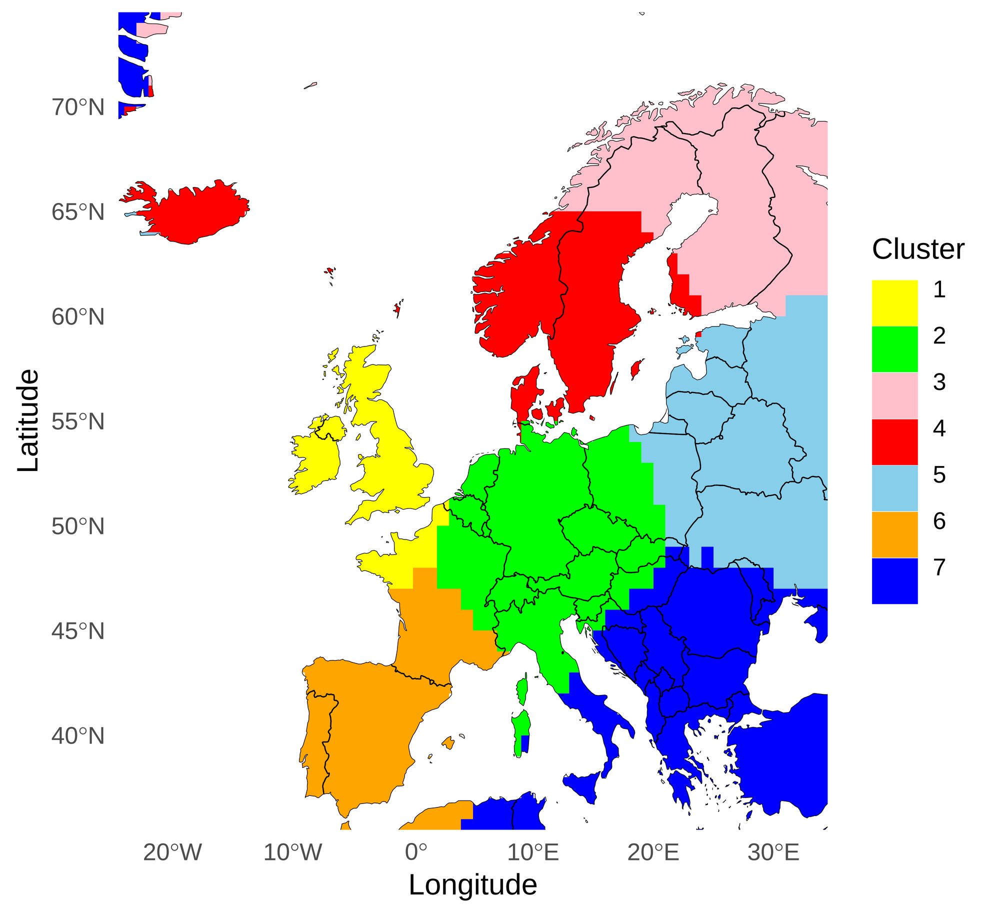


Figure 1. Cluster map of summer (JJA) temperature anomalies from the Berkeley dataset. The seven spatial clusters were identified using the partitioning around medoids (PAM) algorithm applied to a distance based on pairwise tail dependence.

Marginal model

The marginal distribution of each location $i \in \{1,\dots,n_j\}$ in a given cluster $j \in \{1,\dots,7\}$ is modelled with a nonstationary generalized extreme value distribution:

$$G_{i,j}(z) = \exp\left\{-\left[1+\xi_{i,j}\left(\frac{z-\mu_{i,j}(t)}{\sigma_{i,j}}\right)\right]_+^{-1/\xi_{i,j}}\right\}$$

$$\mu_{i,j}(t) = \mu_{0,i,j} + \mu_{1,i,j}\text{GMST}(t) + \mu_{2,i,j}\text{ENSO3.4}(t) + \mu_{3,i,j}\text{NAO}(t)$$

where GMST(t) is the annual Global Mean Surface Temperature anomaly, ENSO3.4(t) is the July-centered 5-month running mean Niño 3.4 anomaly, and NAO(t) is the standardized summer (JJA) North Atlantic Oscillation index.

We define the parameter vector for location i in cluster j as follows:

$$\boldsymbol{\beta}_{i,j} := (\mu_{0,i,j}, \mu_{1,i,j}, \mu_{2,i,j}, \mu_{3,i,j}, \sigma_{i,j}, \xi_{i,j})^\top \in \mathbb{R}^6$$

We penalize variation across neighboring locations. Let $\boldsymbol{\gamma}_{p,j} \in \mathbb{R}^{n_j}$ be the vector of the p-th parameter across all n_j locations in cluster j.

We estimate the parameters by minimizing a penalized negative log-likelihood [5,6,7]:

$$\hat{\boldsymbol{\beta}}_j = \arg\min_{\boldsymbol{\beta}_j}\left[-\log\mathcal{L}(\boldsymbol{\beta}_j) + \sum_{p=1}^{4}\lambda_{j,p}\boldsymbol{\gamma}_{j,p}^\top \boldsymbol{S}_j \boldsymbol{\gamma}_{j,p}\right]$$

where $\boldsymbol{S}_j \in \mathbb{R}^{n_j\times n_j}$ is an adjacency matrix:

$$s_{lk} = \begin{cases} |\mathcal{N}_l| & \text{if } l=k \\ -1 & \text{if } l \text{ and } k \text{ are neighbors} \\ 0 & \text{otherwise} \end{cases}$$

and $\mathcal{N}_l$ denotes the set of neighbors of location l.

Spatio-temporal modeling

We model extremal dependence via the copula of a latent process [8,9] that is applied on all daily temperatures:

$$Z(\mathbf{s},t) = R(t)^\delta W(\mathbf{s},t)^{1-\delta}, \quad \delta\in[0,1]$$

- $R(t)$: IID standard unit-Pareto
- $W(\mathbf{s},t) = \frac{1}{1-\Phi(W^*(\mathbf{s},t))}$: space-time process
- $W^*(\mathbf{s},t)$: zero-mean Gaussian field with separable covariance

$$C((\mathbf{s}_l,t_q),(\mathbf{s}_k,t_r)) = C_s(\mathbf{s}_l,\mathbf{s}_k)C_t(t_q,t_r)$$

Spatial covariance (Spatially nonstationary, anisotropic) [10]:

$$C_s(\mathbf{s}_l,\mathbf{s}_k) = \frac{\left(\sqrt{2}\theta_l^{1/4}\theta_k^{1/4}/(\theta_l+\theta_k)^{1/2}\right)^2}{1+\left(\frac{\|A(\mathbf{s}_l-\mathbf{s}_k)\|}{\theta_l+\theta_k}\right)^2}$$

- $\theta_i \sim$ Gamma(α,τ): random local range
- A: anisotropy matrix

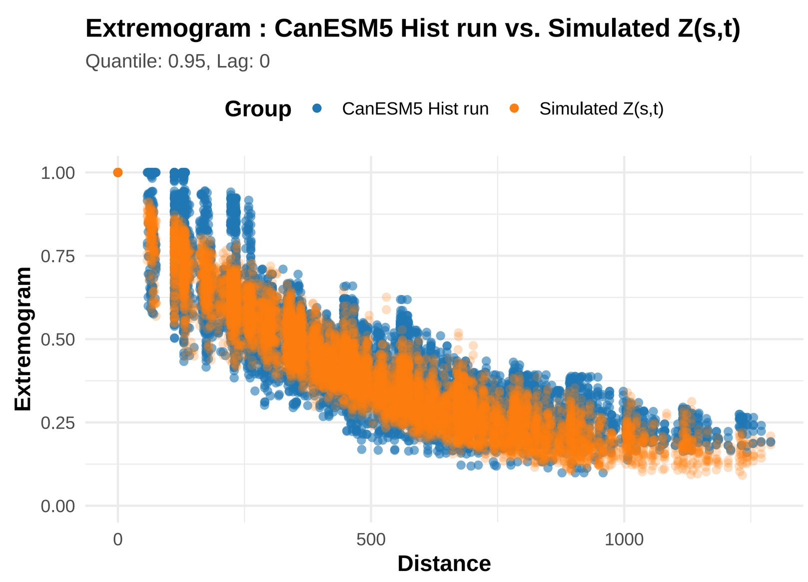
$$A = \begin{pmatrix}\cos\omega & -\sin\omega\\ \sin\omega & \cos\omega\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & \rho\end{pmatrix}$$

Temporal covariance [11]:

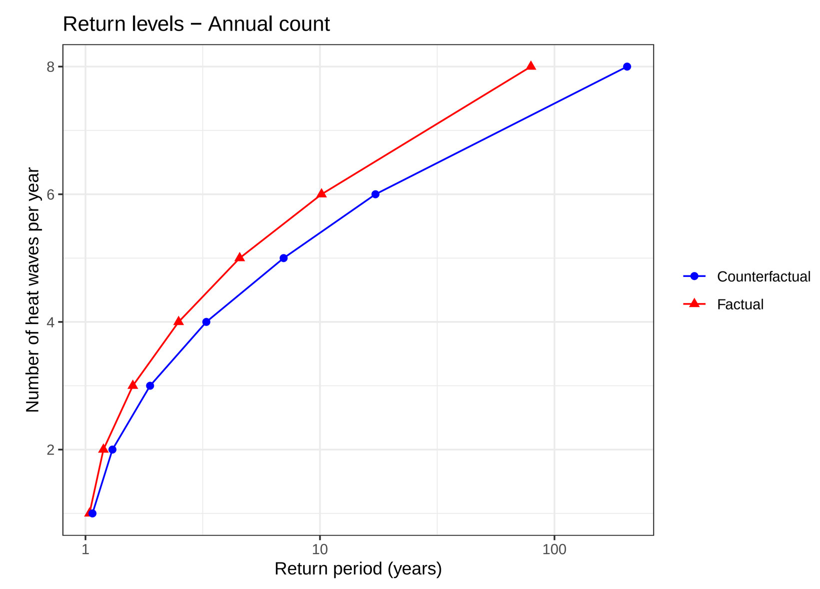
$$C_t(t_q,t_r) = \frac{2^{1-\nu}}{\Gamma(\nu)}\left(\frac{\sqrt{2\nu}\,|t_q-t_r|}{\kappa}\right)^\nu K_\nu\left(\frac{\sqrt{2\nu}\,|t_q-t_r|}{\kappa}\right)$$

- Likelihood-free parameter estimation with neural Bayes estimators [12].

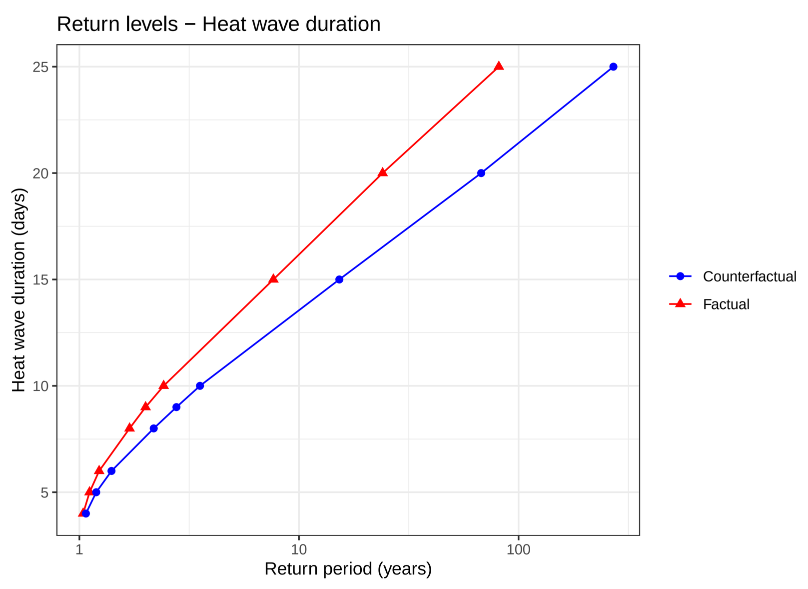
Preliminary results on CanESM5 model [13] cluster 1



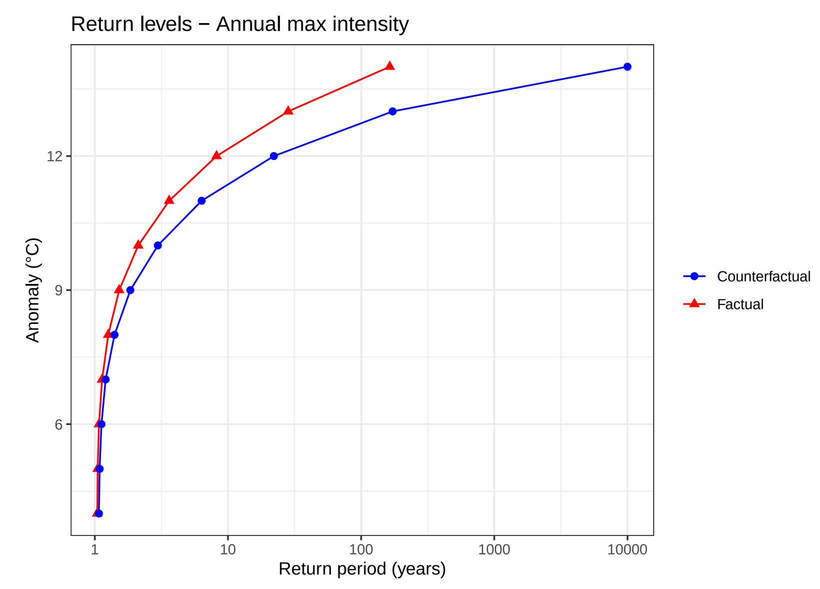
(a) Empirical extremogram at lag 0 as a function of spatial distance Cluster 1. Results from historical data and simulations of the $Z(\mathbf{s},t)$ process are shown for comparison.



(b) Return levels of heat wave number during a summer Cluster 1.



(c) Return levels of heat wave duration Cluster 1.



(d) Return levels of heat wave intensity Cluster 1.

Figure 2. Preliminary results on CanESM5 model cluster 1. Historical (Factual) and natural (Counterfactual) scenarios are shown for comparison.

References

[1] Eyring, V. et al. "Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization", Geoscientific Model Development, 2016.
[2] Rohde, R. A. and Hausfather, Z. "The Berkeley Earth Land/Ocean Temperature Record", Earth System Science Data, 2020.
[3] Bernard, E. et al. "Clustering of maxima: Spatial dependencies among heavy rainfall in France", Journal of Climate, 2013.
[4] Kiriliouk, A. and Naveau, P. "Climate extreme event attribution using multivariate peaks-over-thresholds modeling and counterfactual theory", 2020.
[5] Youngman, B. D. "evgam: An R package for generalized additive extreme value models", Journal of Statistical Software, 2022.
[6] Robben, J. et al. "The short-term association between environmental variables and mortality: evidence from Europe", arXiv preprint arXiv:2405.18020, 2024.
[7] Auld, G. et al. "Changes in the distribution of annual maximum temperatures in Europe", Advances in Statistical Climatology, Meteorology and Oceanography, 2023.
[8] Dell'Oro, L. and Gaetan, C. "Flexible space-time models for extreme data", arXiv preprint arXiv:2411.19184, 2024.
[9] Huser, R. and Wadsworth, J. L. "Modeling spatial processes with unknown extremal dependence class", Journal of the American Statistical Association, 2019.
[10] Coube-Sisqueille, S. et al. "Nonstationary Spatial Process Models with Spatially Varying Covariance Kernels", arXiv preprint arXiv:2203.11873, 2022.
[11] Stein, M. L. "Interpolation of Spatial Data: Some Theory for Kriging", Springer, 1999.
[12] Sainsbury-Dale, M. et al. "Likelihood-free parameter estimation with neural Bayes estimators", The American Statistician, 2024.
[13] Swart, N. C. et al. "The Canadian Earth System Model version 5 (CanESM5.0.3)", Geoscientific Model Development, 2019.

Acknowledgements

This project ("EXALT") was carried out with the financial support of the "Actions de recherche concertées (ARC)" programme of the Wallonia-Brussels Federation. Computational resources have been provided by the Consortium des Équipements de Calcul Intensif (CÉCI), funded by the Fonds de la Recherche Scientifique de Belgique (F.R.S.-FNRS) under Grant No. 2.5020.11 and by the Walloon Region